UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/685,656	10/14/2003	W. Todd Daniell	030220; 190250-1300	5661

38823   7590   06/12/2009
AT&T Legal Department - TKHR
Attn: Patent Docketing
One AT&T Way
Room 2A-207
Bedminster, NJ 07921

EXAMINER
MACILWINEN, JOHN MOORE JAIN

ART UNIT	PAPER NUMBER
2442	

MAIL DATE	DELIVERY MODE
06/12/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

	Application No.	Applicant(s)
	10/685,656	DANIELL ET AL.
Office Action Summary	Examiner	Art Unit	
	John M. MacIlwinen	2442	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *04 May 2009*.
2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.
3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is
   closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,6,11-14,16,17 and 19-39* is/are pending in the application.
   4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5)☐ Claim(s) _____ is/are allowed.
6)☒ Claim(s) *1,6,11-14,16,17 and 19-39* is/are rejected.
7)☐ Claim(s) _____ is/are objected to.
8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.
10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.
   Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
   Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a)☐ All   b)☐ Some *   c)☐ None of:
      1.☐ Certified copies of the priority documents have been received.
      2.☐ Certified copies of the priority documents have been received in Application No. _____ .
      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage
         application from the International Bureau (PCT Rule 17.2(a)).
   * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
   Paper No(s)/Mail Date *5/12/2009*.

4) ☐ Interview Summary (PTO-413)
   Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____ .

DETAILED ACTION

Response to Arguments

1.   Applicant's arguments filed 5/4/2009 have been fully considered.

2.   Applicant begins, on page 15, arguing that "While the Office Action argues that a combination of eight (8) references renders claim 1 obvious, Applicants disagree". In response to applicant's argument that the examiner has combined an excessive number of references, reliance on a large number of references in a rejection does not, without more, weigh against the obviousness of the claimed invention. See *In re Gorman*, 933 F.2d 982, 18 USPQ2d 1885 (Fed. Cir. 1991).

3.   Applicant's remaining arguments, pages 16 - 38, fail to comply with 37 CFR 1.111(b) because they amount to a general allegation that the claims define a patentable invention without specifically pointing out how the language of the claims patentably distinguishes them from the references. Applicant's arguments merely recite the amended claim language and state that individual references "fail to even suggest" said claim language.

4.   Regarding claim 1, Applicant's arguments are additionally unpersuasive because Gordon (US 6,732,157 B1) teaches the newly added limitations including searching for the non-displaying characters in the email and removing the searched non-displaying characters (Gordon, col. 9 lines 50 - 55, showing "remov[ing] various formatting specific to the protocols associated with the electronic mail messages" prior to further processing). As Gordon shows searching for and removing non-displaying characters in

an email message, said message inherently comprised both displaying and non-

displaying characters.

5.   However, Applicant's amendments to claims 6, 23, 24, 25, 30, 31 and 32 have

necessitated the new grounds of rejection, discussed further below.

Specification

6.   The specification is objected to as failing to provide proper antecedent basis for

the claimed subject matter for the reasons given below in the 35 USC 112 written

description rejection. See 37 CFR 1.75(d)(1) and MPEP § 608.01(o).

Claim Rejections - 35 USC § 112

7.   The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of
> making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the
> art to which it pertains, or with which it is most nearly connected, to make and use the same and shall
> set forth the best mode contemplated by the inventor of carrying out his invention.

8.   Claims 1, 2, 23, 24, 25, 30, 31 and 32 are rejected under 35 U.S.C. 112, first

paragraph, as failing to comply with the written description requirement. The claim(s)

contains subject matter which was not described in the specification in such a way as to

reasonably convey to one skilled in the relevant art that the inventor(s), at the time the

application was filed, had possession of the claimed invention. Said claims recite

"searching for" or "searching logic configured to search for" the "non-displaying

characters" as well as "removing the searched non-displaying characters". "searching

for" and "searching logic configured to search for" and "removing the searched" lack

written description.

9.   Additionally, claim 6 recites "the displaying characters of the STMP email

address"; there is a lack written description for said limitation.

10.   The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly
> claiming the subject matter which the applicant regards as his invention.

11.   Claims 6, 23, 24, 25, 30, 31 and 32 are rejected under 35 U.S.C. 112, second

paragraph, as being indefinite for failing to particularly point out and distinctly claim the

subject matter which applicant regards as the invention.

12.   Regarding claim 6, said claim recites "a token representative of the displaying

characters of the SMTP email address". It is unclear what "the displaying characters of

the STMP email address" refers to.

13.   Regarding claims 23, 24, 25, said claims refer to "an SMTP email address . . .

and an address". It is unclear what said "and an address" refers to.

14.   Regarding claims 30, 31 and 32, said claims recite "tokenize logic configured to

tokenize the attachment"; however, said claims also recite "generated tokens"; the claim

language only recites a singular token and thus it is unclear what said "tokens" refers to.

Said claims also recite "where only the displaying characters are tokenized"; however,

said claims only recite tokenizing the attachment, and thus it is unclear what the relation

is between "displaying characters" and "the attachment"; that is, how there are

"displaying characters" relating to the tokenization of "then attachment".

15.   Claims 23, 24, 25, 30, 31, 32, 33 and 35 recites the limitation "the attachment".

There is insufficient antecedent basis for this limitation in the claims.

16.   In order to perform a complete examination, the above claims have been

interpreted broadly.

Claim Rejections - 35 USC § 101

17.   35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of
> matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the
> conditions and requirements of this title.

18.   Claims 23, 24, 25 - 29, 30 - 38 are rejected under 35 U.S.C. 101 because the

claimed invention is directed to non-statutory subject matter.

19.   Regarding claim 23, said claim recites a system comprising email receive logic,

searching logic, removing logic, tokenize logic, analysis logic and sorting logic.

Applicant's specification, pg. 8, describes an email application having a "filter 220" and

pg. 9, describes the "filter 220" as comprising said "logic". Thus, claim 23 is solely

directed to nonstatutory subject matter, i.e., software.

20.   Regarding claims 25 – 29 and 32 – 38, said claims are directed to "a computer-

readable medium". However, Applicant's specification on pg. 22 states that "a

computer-readable medium can be . . . an electronic, magnetic, optical, infrared . . . or

propagation medium . . . [or] an optical fiber . . ."; thus, said claims are directed to

nonstatutory subject matter.

21.   Regarding claims 24 and 31, said claims recite a "system comprising means for"

where each of the means for, based on Applicant's specification, appears directed to

software (see the above discussion of claim 23 above). It thus appears that Applicant's

system consists of software and thus is directed to non-statutory subject matter.

22.   Regarding claim 30, said claim is directed to "a memory component" containing

various "logic". Said claim thus appears to be directed to the same media as claims 25-

29 and 32 - 38 and thus is also directed to non-statutory subject matter for the reasons

given above.

Claim Rejections - 35 USC § 103

23.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

24.   Claims 1 and 39 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Shipp (US 2004/0093384 A1) in view of Devine et al. (US 6,968,571 B2), hereafter

Devine, further in view of Milliken et al. (US 2004/0073617 A1), hereafter Milliken,

further in view of Anderson et al. (US 2004/0064537 A1), hereafter Anderson, further in

view of Uuencode and MIME FAQ

(http://web.archive.org/web/20021217052047/http://users.rcn.com/wussery/attach.html),

further in view of Gordon et al. (US 6,732,157 B1), hereafter Gordon, further in view of

Sahami et al. (A Bayesian Approach to Filtering Junk E-Mail), hereafter Sahami, further

in view of Woitaszek and Shaaban (Identifying Junk Electronic Mail in Microsoft Outlook

with a Support Vector Machine), hereafter Woitaszek, and Burdick (US 2004/0107189

A1).

       Regarding claim 1, Shipp shows a method comprising

receiving an email message from a simple mail transfer protocol (SMTP) server,

the email message comprising ([0018,0023]) displaying characters ([64-67]) and non-

displaying characters ([57-58, 61-73]), the email message further comprising

a text body ([0064,0065])

an SMTP email address ([0018,0023,0039,0045,0046])

a domain name corresponding to the SMTP email address ([0039,0045,0046])

an attachment ([0081])

tokenizing the text body to generate tokens representative of words in the text

([0064-0067])

tokenizing the SMTP email address to generate a token representative of the

SMTP email address ([0039,0043,0069])

tokenizing the domain name to generate a token that is representative domain

name ([0022])

as well as showing MD5 hashing ([0093]).

Shipp does not show a 32-bit string indicative of the length of the email message,

nor does Shipp show searching for and removing the non-displaying characters in the

email, determining and filtering the non-alphabetic displaying characters in the email,

generating a phonetic equivalent for each word that includes only alphabetic displaying

characters that has a phonetic equivalent, tokenizing the attachment and the steps

comprising tokenizing said attachment, determining a probability value for each

generated token, selecting a predefined number of interesting tokens, the interesting

tokens being the generated tokens having the greatest non-neutral probability value;

performing a Bayesian analysis on the selected interesting tokens to generate a spam probability; and categorizing the email message as a function of the generated spam probability.

Devine shows utilizing a 32-bit string in a message header which is indicative of the total length of said message (col. 24 lines 52-67).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp with that of Devine in order to better identify message contents so as to facilitate leveraging common code for processing messages (Devine col. 23 lines 60-61).

Shipp in view of Devine do not show tokenizing the attachment.

Milliken shows tokenizing the attachment to generate a token that is representative of the attachment, the tokenizing steps comprising the steps of generating a MD5 hash of the attachment ([0010-0013 and 0052]where MD5 hashes are inherently 128-bit).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Devine with that of Milliken in order to better identify spam email, as at the time of Shipp's disclosure, spam email was thought "currently" not to be associated with attachments ([81]), an area for which Milliken's more recent disclosure provides updated guidance.

Shipp in view of Devine and Milliken do not show appending the 32-bit string to the generated MD5 hash to produce a 160-bit number.

Anderson shows ([0057-0059]) appending an MD5 hash (inherently 128-bits) to network transmission size information (shown by Devine to be said 32-bit string, and where 32 +128 is inherently 160).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Devine and Milliken with that of Anderson in order to better uniquely identify messages (Anderson [0057-0059]), leading to improved message spam identification.

Shipp in view of Devine, Milliken and Anderson do not show UUencoding said 160-bit number to generate a token representative of the attachment.

Uuencode and MIME FAQ shows UUencoding a file.

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Devine, Milliken and Anderson with that of Uuencode and MIME FAQ in order to store the message identification information (represented by the 160-bit number shown by Shipp in view of Devine, Milliken and Anderson) in a format easily exchanged over email (UUencode and MIME FAQ) since UUencoding produces an easily emailed file and since the disclosure of Shipp in view of Devine, Milliken and Anderson relates to email and files transferred over email. Furthermore, UUencoding is a prior art element, as shown in UUencode and MIME FAQ, and thus UUencoding the 160-bit number is combing a prior art element (UUencoding) to known methods (the known methods shown by Shipp in view of Devine, Milliken and Anderson) to yield predictable results (the results being a UUencoded item).

Shipp in view of Devine, Milliken, Anderson and UUencode and MIME FAQ do

not show determining a probability value for each of the generated tokens.

Gordon shows determining a probability value for each of the generated tokens

(col. 11 lines 15 –55) along with, as an initial processing step, searching for non-

displaying characters in the email and removing the non-displaying characters in the

email (col. 9 lines 55 - 60, Fig. 7).

It would have been obvious to one of ordinary skill in the art at the time of the

invention to modify the disclosure of Shipp in view of Devine, Milliken, Anderson and

Uuencode and MIME FAQ with that of Gordon in order to better identify spam elements

in messages (Gordon col. 11 lines 15 –55).

Shipp in view of Devine, Milliken, Anderson, UUencode and MIME FAQ and

Gordon do not show sorting the generated tokens in accordance with the corresponding

determined spam probability value to determine a predefined number of interesting

tokens, the predefined number of interesting tokens being a subset of the generated

tokens, selecting the predefined number of interesting tokens, the interesting tokens

being the generated tokens having the greatest non-neutral probability value;

performing a Bayesian analysis on the selected interesting tokens to generate a spam

probability; and categorizing the email message as a function of the generated spam

probability.

Sahami shows selecting a predefined number of interesting tokens, the

interesting tokens being the generated tokens having the greatest non-neutral

probability value to determine a predefined number of interesting tokens, the predefined

number of interesting tokens being a subset of the generated tokens (pg. 4, col. 1,
showing having initially "several thousand" features, then selecting 500 of said features
after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and then
selecting, of the remaining feature, the 500 features with the highest non-neutral
probability value (pg. 6, col. 1, paragraph 1)); performing a Bayesian analysis on the
selected interesting tokens to generate a spam probability; and categorizing the email
message as a function of the generated spam probability (pg. 2, col. 2; pg. 4, col. 2; pg.
6, col. 1).

It would have been obvious to one of ordinary skill in the art at the time of the
invention to modify the disclosure of Shipp in view of Devine, Milliken, Anderson,
Uuencode and MIME FAQ and Gordon with that of Sahami in order to more accurately
identify spam email.

Shipp in view of Devine, Milliken, Anderson, UUencode and MIME FAQ and
Gordon and Sahami thus do show selecting a subset of the generated tokens based on
probability value as well as where the interesting tokens are a subset of the generated
tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show explicitly show where the
tokens are sorted in accordance with the corresponding determined spam probability
value.

Woitaszek shows where the tokens are sorted in accordance with the
corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the
invention to modify the disclosure of Shipp in view of Devine, Milliken, Anderson,

Uuencode and MIME FAQ, Gordon and Sahami with that of Woitaszek in order to
arrange the calculated values in a logical manner, enabling a simple method of
extracting the most interesting results (Sahami's disclosure involving selecting said
most interesting tokens) via simply taking the top occurring results in Woitaszek's sorted
list, as well as to include the abilities to integrate the spam software into a commonly
used email program (Woitaszek, Abstract, pg. 1 col. 2).

Shipp in view of Devine, Milliken, Anderson, Uuencode and MIME FAQ, Gordon,
Sahami and Woitaszek do not explicitly show determining the non-alphabetic displaying
characters in the email, filtering the determined non-alphabetic displaying characters
from the email, and generating a phonetic equivalent for each word that includes only
alphabetic displaying characters that has a phonetic equivalent.

Burdick shows determining the non-alphabetic displaying characters in the email,
filtering the determined non-alphabetic displaying characters from the email, and
generating a phonetic equivalent for each word that includes only alphabetic displaying
characters that has a phonetic equivalent ([14]).

It would have been obvious to one of ordinary skill in the art at the time of the
invention to modify the disclosure of Shipp in view of Devine, Milliken, Anderson,
Uuencode and MIME FAQ, Gordon and Sahami and Woitaszek with that of Burdick in
order to ensure the data (that is, email message contents) is in good form before it is
further processed, thus increasing the ease of using the data and its utility (Burdick, [2-
4]).

Shipp in view of Devine, Milliken, Anderson, Uuencode and MIME FAQ, Gordon,

Sahami, Woitaszek and Burdick thus show claim 1.

25.   Regarding claim 39, Shipp in view of Devine, Milliken, Anderson, Uuencode and

MIME FAQ, Gordon, Sahami, Woitaszek and Burdick show wherein the email is

received at a computing device (Milliken, Abstract, Shipp, Abstract).

26.   Claims 6, 11, 12, 13, 14, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30,

31, 32, 33, 34, 35, 36, 37 and 38 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Shipp in view of Milliken, Sahami, Woitaszek and Gordon.

27.   Regarding claim, 6 Shipp shows a method comprising receiving an email

message comprising a text body ([64,65]), an SMTP email address ([39,43,69]), and a

domain name corresponding to the SMTP email address ([39,45,46]), the text body

including displaying characters ([64-67]) and non-displaying characters ([57-58, 61-73]);

tokenizing the SMTP email address to generate a token representative of the

SMTP email address ([39,43,63])

tokenizing the domain name to generate a token representative of the domain

name ([22]), and determining a spam probability value from the generated tokens

([14,76]).

Shipp does not show tokenizing the attachment to generate a token that is

representative of the attachment.

Milliken shows tokenizing the attachment to generate a token that is

representative of the attachment ([10-13 and 51 – 53]).

It would have been obvious to one of ordinary skill in the art at the time of the

invention to modify the disclosure of Shipp with that of Milliken in order to better identify

spam email, as at the time of Shipp's disclosure, spam email was thought "currently" not

to be associated with attachments ([81]); spam and attachments are however an area

for which Milliken's more recent disclosure provides updated guidance.

Shipp in view of Milliken do not show explicitly show where the tokens are sorted

in accordance with the corresponding determined spam probability value to determine a

predefined number of interesting tokens, the predefined number of interesting tokens

being a subset of the generated tokens.

Sahami shows selecting a predefined number of interesting tokens, the

predefined number of interesting tokens being a subset of the generated tokens (pg. 4,

col. 1, showing having initially "several thousand" features, then selecting 500 of said

features after first sorting out features that occur fewer than 3 times (pg. 4, col. 2) and

then selecting, of the remaining feature, the 500 features with the highest non-neutral

probability value (pg. 6, col. 1, paragraph 1)).

It would have been obvious to one of ordinary skill in the art at the time of the

invention to modify the disclosure of Shipp in view of Milliken with that of Sahami in

order to more accurately identify spam email (Sahami, Abstract).

Shipp in view Milliken and Sahami thus do show selecting a subset of the

generated tokens based on probability value as well as where the interesting tokens are

a subset of the generated tokens (Sahami, pg. 6, col. 1, paragraph 1), but do not show

explicitly show where the tokens are sorted in accordance with the corresponding

determined spam probability value.

Woitaszek shows where the tokens are sorted in accordance with the corresponding determined spam probability value (Tables 4 and 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken and Sahami with that of Woitaszek in order to arrange the calculated values in a logical manner, enabling a simple method of extracting the most interesting results (as discussed by Sahami) via simply taking the top occurring results in Woitaszek's sorted list, as well as to include the abilities to integrate the spam software into a commonly used email program (Woitaszek, Abstract, pg. 1 col. 2).

Shipp in view of Milliken, Sahami and Woitaszek do not show searching for non-displaying characters in the email and removing the non-displaying characters in the email.

Gordon shows, as an initial processing step, searching for non-displaying characters in the email and removing the non-displaying characters in the email (col. 9 lines 55 - 60, Fig. 7).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify the disclosure of Shipp in view of Milliken, Sahami and Woitaszek with that of Gordon in order to prepare a version of the data more amicable to future processing (Gordon, col. 9 line 56).

28.   Regarding claim 16, Shipp in view of Milliken, Sahami, Woitaszek and Gordon further show receiving an email message including a text body (Shipp [64,65]).

29.   Regarding claim 17, Shipp in view of Milliken, Sahami, Woitaszek and Gordon

further show tokenizing the words in the text body to generate tokens representative of

the words in the text body (Shipp [64,65]).

30.   Regarding claim 23, Shipp in view of Milliken, Woitaszek and Gordon further

show a system comprising a text body (Shipp, [64,65]), an SMTP email address (Shipp,

[39-43,69]), and a domain name corresponding to the SMTP email address and an

address (Shipp, [39,45,46]) the email message further including (Shipp, [64-67]) and

non-displaying characters (Shipp, [57-58, 61-73]);

        searching logic configured to search for the non-displaying characters in the

email;

        removing logic configured to remove the searched non-displaying characters

(Gordon, col. 9 lines 55 - 60)

        tokenizing logic configured to tokenize the SMTP email address to generate a

token representative of the SMTP email address (Shipp, [39,43,63])

        tokenizing logic configured to tokenize the attachment to generate a token that is

representative of the attachment (Milliken [10-13 and 51 – 53])

        tokenizing the domain name to generate a token representative of the domain

name (Shipp, [22]), and

        determining a spam probability value from the generated tokens (Shipp, [14,76])

and

        sorting logic configured to sort the generated tokens in accordance with the

corresponding determined spam probability value (Woitaszek, Abstract, Tables 4 and 5)

to determine a predefined number of interesting tokens, the predefined number of

interesting tokens being a subset of the generated tokens (Sahami, pgs. 4 and 6)

wherein only displaying characters are tokenized (Gordon, col. 9 lines 55 – col.

10 line 19, Fig. 17).

31.   Regarding claim 24, Shipp in view of Milliken, Sahami, Woitaszek and Gordon

further show means for receiving an SMTP email address, and a domain name

corresponding to the SMTP email address (Shipp, [39,45,46]) and an address (Shipp,

[39,45,46]) the email message further including (Shipp, [64-67]) and non-displaying

characters (Shipp, [57-58, 61-73]);

means for searching for the non-displaying characters in the email;

means for removing the searched non-displaying characters (Gordon, col. 9 lines

55 - 60)

means for tokenizing the SMTP email address to generate a token representative

of the SMTP email address (Shipp, [39,43,63])

means for tokenizing the attachment to generate a token that is representative of

the attachment (Milliken [10-13 and 51 – 53])

means for tokenizing the domain name to generate a token representative of the

domain name (Shipp, [22]), and

means for determining a spam probability value from the generated tokens

(Shipp, [14,76]) and

sorting logic configured to sort the generated tokens in accordance with the

corresponding determined spam probability value (Woitaszek, Abstract, Tables 4 and 5)

to determine a predefined number of interesting tokens, the predefined number of

interesting tokens being a subset of the generated tokens (Sahami, pgs. 4 and 6)

wherein only displaying characters are tokenized (Gordon, col. 9 line 55 - col. 10 line

19, Fig. 7).

32.   Regarding claim 25, Shipp in view of Milliken, Sahami, Woitaszek and Gordon

further show a computer-readable medium that includes a program, that when executed

by a computer, performs the actions of receive an email message comprising an SMTP

email address, ([39-43,69]), a domain name corresponding to the SMTP email address

([39,0045,46]) and an address (Shipp, [39,45,46]) the email message further including

(Shipp, [64-67]) and non-displaying characters (Shipp, [57-58, 61-73]);

        search for the non-displaying characters in the email;

        remove the searched non-displaying characters (Gordon, col. 9 lines 55 - 60)

        tokenizing the SMTP email address to generate a token representative of the

SMTP email address (Shipp, [39,43,63])

        tokenizing the attachment to generate a token that is representative of the

attachment (Milliken [10-13 and 51 – 53])

        tokenizing the domain name to generate a token representative of the domain

name ([0022]), and

        determining a spam probability value from the generated tokens ([0014,0076])

and

        sorting logic configured to sort the generated tokens in accordance with the

corresponding determined spam probability value (Woitaszek, Abstract, Tables 4 and 5)

to determine a predefined number of interesting tokens, the predefined number of

interesting tokens being a subset of the generated tokens (Sahami, pgs. 4 and 6)

wherein only displaying characters are tokenized (Gordon, col. 9 line 55 - col. 10 line

19, Fig. 7).

33.   Regarding claim 30, Shipp in view of Milliken, Sahami, Woitaszek and Gordon

further show a system comprising a memory component that stores email logic

configured to receive an email message comprising an address (Shipp, [39,45,46]) the

email message further including (Shipp, [64-67]) and non-displaying characters (Shipp,

[57-58, 61-73]);

 search logic configured to search for the non-displaying characters in the email;

 remove logic configured to remove the searched non-displaying characters

(Gordon, col. 9 lines 55 - 60)

 tokenize logic configured to tokenize the entire attachment to generate a token

representative of the attachment (Milliken [10-13 and 70]); and

 analysis logic configured to determine a spam probability values from the

generated tokens (Milliken [10-13] and Shipp [14,76]) and

 sorting logic configured to sort the generated tokens in accordance with the

corresponding determined spam probability value (Woitaszek, Abstract, Tables 4 and 5)

to determine a predefined number of interesting tokens, the predefined number of

interesting tokens being a subset of the generated tokens (Sahami, pgs. 4 and 6)

wherein only displaying characters are tokenized (Gordon, col. 9 line 55 - col. 10 line

19, Fig. 7).

34.   Regarding claim 31, Shipp in view of Milliken, Woitaszek and Gordon further

show means for receiving an email message comprising (Shipp [18,23]) an address

(Shipp, [39,45,46]) the email message further including (Shipp, [64-67]) and non-

displaying characters (Shipp, [57-58, 61-73]);

   means for searching for the non-displaying characters in the email;

   means for removing the searched non-displaying characters (Gordon, col. 9 lines

55 - 60)

   means for tokenizing the attachment to generate a token representative of the

attachment (Milliken [10-13 and 70]); and

   means for determining a spam probability values from the generated tokens

(Milliken [10-13] and Shipp [14,76])

   sorting logic configured to sort the generated tokens in accordance with the

corresponding determined spam probability value (Woitaszek, Abstract, Tables 4 and 5)

to determine a predefined number of interesting tokens, the predefined number of

interesting tokens being a subset of the generated tokens (Sahami, pgs. 4 and 6)

wherein only displaying characters are tokenized (Gordon, col. 9 line 55 - col. 10 line

19, Fig. 7).

35.   Regarding claim 32, Shipp in view of Milliken, Sahami, Woitaszek and Gordon

further shows a computer-readable medium that when executed by a computer,

performs at least the following: receive an email message comprising an attachment

(Shipp [18,23] and Milliken [10-13])),

   tokenize logic configured to tokenize the entire attachment to generate a token

representative of the attachment (Milliken [10-13 and 70]); and

determine a spam probability values from the generated tokens (Milliken [10-13] and

Shipp [14,76]) and

    sort the generated tokens in accordance with the corresponding determined

spam probability value (Woitaszek, Abstract, Tables 4 and 5) to determine a predefined

number of interesting tokens, the predefined number of interesting tokens being a

subset of the generated tokens (Sahami, pgs. 4 and 6).

36.   Regarding claims 11 and 26, Shipp in view of Milliken, Sahami, Woitaszek and

Gordon show assigning a spam probability value to the token representative of the

SMTP email address (Shipp [18,23,39,40-43], Woitaszek, Tables 4 and 5) and

    assigning a spam probability value to the token representative of the domain

name (Shipp [22]).

    and generating a Bayesian probability values using the spam probability values

assigned to the tokens (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

37.   Regarding claims 12 and 27 Shipp in view of Milliken, Sahami, Woitaszek and

Gordon further show comparing the generated Bayesian probability value with a

predefined threshold value (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

38.   Regarding claims 13 and 28 Shipp in view of Milliken, Sahami, Woitaszek and

Gordon further show categorizing the email message as spam in response to the

Bayesian probability value being greater than the predefined threshold (Sahami, pg.2,

col. 2; pg. 4, col. 2; pg. 6, col. 1).

39.   Regarding claims 14 and 29 Shipp in view of Milliken, Sahami, Woitaszek and Gordon further show categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold (Sahami pg. 6 col. 1).

40.   Regarding claims 19 and 35, Shipp in view of Milliken, Sahami, Woitaszek and Gordon show assigning a spam probability value to each of the tokens representation of the words in the text body (Woitaszek, Tables 4 and 5)

assigning a spam probability value to token representative of the attachment (Woitaszek, Tables 4 and 5, and Milliken, [10-13]),

and generating a Bayesian probability value using the spam probability values assigned to the token (Sahami, pg. 4 col. 2).

41.   Regarding claims 20 and 36, Shipp in view of Milliken, Sahami, Woitaszek and Gordon further show comparing the generated Bayesian probability value with a predefined threshold value (Sahami, pg. 4 col. 2).

42.   Regarding claims 21 and 37, Shipp in view of Milliken, Sahami, Woitaszek and Gordon further show categorizing the email message as spam in response to the Bayesian probability value being greater than the predefined threshold (Sahami, pg.2, col. 2; pg. 4, col. 2; pg. 6, col. 1).

43.   Regarding claims 22 and 38, Shipp in view of Milliken, Sahami, Woitaszek and Gordon further show categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold (Sahami, pg. 6 col. 1).

44.   Regarding claim 33, Shipp in view of Milliken, Sahami, Woitaszek and Gordon

further show receiving an email message including a text body (Shipp [64,65]).

45.   Regarding claim 34, Shipp in view of Milliken, Sahami, Woitaszek and Gordon

further show tokenizing the words in the text body to generate tokens representative of

the words in the text body (Shipp [64,65]).

Conclusion

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to John M. MacIlwinen whose telephone number is (571)

272-9686. The examiner can normally be reached on M-F 7:30AM - 5:00PM EST; off

alternate Fridays.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Andrew Caldwell can be reached on (571) 272-3868. The fax phone

number for the organization where this application or proceeding is assigned is 571-

273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Andrew Caldwell/
Supervisory Patent Examiner, Art
Unit 2442

John MacIlwinen

(571) 272 - 9686